Suitability of Density Functional Methods for Calculation of Electrostatic Properties

ROBERT SOLIVA,[1] MODESTO OROZCO,[1] F. JAVIER LUQUE[2]

[1] *Department de Bioquímica i Biología Molecular, Facultat de Química, Universitat de Barcelona, Martí i Franquè 1, Barcelona 08028, Spain*

[2] *Departament de Farmàcia, Unitat Fisicoquímica, Facultat de Farmàcia, Universitat de Barcelona, Avgda Diagonal s / n, Barcelona 08028, Spain*

Received 19 August 1996; accepted 31 October 1996

ABSTRACT: A systematic analysis was performed on the suitability of the molecular electrostatic potential (MEP) and MEP-derived properties determined by means of density functional (DFT) methods. Attention was paid to the electrostatic potential (ESP) derived charges, the ESP and exact quantum mechanical dipole moments, the depth of MEP minima, and the MEP distribution in layers around the molecule for a large series of molecules. The electrostatic properties were determined at either local or nonlocal DFT levels using different functionals. The results were compared with the values estimated from quantum mechanical calculations performed at Hartree–Fock, Møller–Plesset up to fourth order, and CIPSI levels. The suitability of the MEP-derived properties estimated from DFT methods is discussed for application in different areas of chemical interest. © 1997 by John Wiley & Sons, Inc. *J Comput Chem* **18:** 980–991, 1997

Keywords: molecular electrostatic potential (MEP); density functional method (DFT); electrostatic potential (ESP) derived charges

Introduction

In the context of quantum mechanics the molecular electrostatic potential (MEP)[1] is defined as the expectation value of the operator $\mathbf{r}^{-1}$ [eq. (1)]. Alternatively, the MEP provides a measure of the electrostatic interaction energy between the molecular charge distribution (electrons and nuclei) and the positive unit charge, which is given by eq. (2) within the molecular orbital linear combination of atomic orbital (MO-LCAO) quantum mechanical (QM) framework.

$$V(\mathbf{r}_1) = \langle \Phi | \hat{r}^{-1} | \Phi \rangle, \qquad (1)$$

Correspondence to: M. Orozco or F. J. Luque.

Contract grant sponsor: DGICYT; contract grant numbers PB93-0779 and PB94-0940

© 1997 by John Wiley & Sons, Inc.

CCC 0192-8651 / 97 / 080980-12

where Φ stands for the unperturbed molecular wave function.

$$V(\mathbf{r}_1) = \sum_A \frac{Z_A}{|\mathbf{r}_1 - \mathbf{R}_A|} - \sum_\mu \sum_\nu P_{\mu\nu} \int \frac{\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})}{|\mathbf{r}_1 - \mathbf{r}|} d\mathbf{r}, \quad (2)$$

where Z_A denotes the nuclear charge of atom A, which is placed at $\mathbf{R}_A$; $P_{\mu\nu}$ is the first-order density matrix; and χ stands for the basis of AO.

The MEP provides a fast, accurate picture of the electrostatic distribution generated from the charge density in the space surrounding the molecule. Owing to the fine details provided in the description of such an electrostatic distribution, the MEP is widely used in different research fields, which include the study of molecular reactivity patterns, noncovalent complexes and biological interactions,[2–12] analysis of molecular similarity,[13–18] description of the crystalline state,[2,19–21] solvation phenomena,[22–27] and topographical analysis of the electronic structure of complex molecules.[28–30] Moreover, the MEP has become very popular as a source of parameters for classical calculations,[31–37] particularly for derivation of atomic partial (electrostatic potential, ESP, derived) charges (Q_N) used in force fields to represent the molecular charge distribution [eq. (3)].

$$\left\{ \mathrm{ABS}\left(\sum_N \frac{Q_\mathrm{N}}{|\mathbf{r}_k - \mathbf{R}_\mathrm{N}|} - \sum_A \frac{Z_A}{|\mathbf{r}_k - \mathbf{R}_A|} + \sum_\mu \sum_\nu P_{\mu\nu} \int \frac{\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})}{|\mathbf{r}_k - \mathbf{r}|} d\mathbf{r} \right) \right\}_k = 0, \quad (3)$$

where $\{\ \}_k$ stands for the set of points considered in the fitting between QM and classical ESP, and N is the number of point charges used to describe the molecular charge distribution.

Most studies in the past examined the dependence of the MEP on the quality of the wave function and the basis set, which influences the elements $P_{\mu\nu}$ of the first-order density matrix and calculation of the monoelectronic integrals in eq. (2). It was generally found that a Hartree–Fock (HF) wave function and a split-valence basis set with polarization functions, like Pople's 6–31G* basis set,[38] are reliable enough to describe the main features of the MEP. Application of this computational level is, however, difficult for large molecules, such as compounds of interest in biological or pharmacological studies. The difficulties are even greater when electron correlation effects have to be considered, which is most cases precludes calculation of the MEP. In these cases the use of semiempirical methods to determine the MEP constitutes an interesting strategy,[39–43] even though caution is needed because notable errors in the details of the MEP may appear. Recently methods based on density functional theory (DFT) have emerged as promising alternatives for calculation of molecular structure and properties, because at the lower levels of theory they are not much more expensive than semiempirical methods while at the higher levels correlation effects are included at a fraction of the computational cost required in standard perturbative or variational methods.

In this study we present the results of a systematic analysis about the quality of the MEP and MEP-derived properties computed from either local and nonlocal DFT using different functionals. Such an analysis is performed by comparison of the results determined from DFT calculations with regard to those derived from QM computations at the HF, Møller–Plesset[44] (MP) theory up to fourth order, and CIPSI[45] levels. The purpose of this study is to examine the use of DFT methods as a practical alternative to MPx calculations to introduce electron correlation effects in the MEP, and the suitability of DFT methods as a source for derivation of the MEP and MEP-derived properties.

Methods

All molecules included in the study were optimized at the HF/6–31G* level using the standard energy minimization procedure in Gaussian 94.[46] These geometries were then used to perform single-point calculations at the different levels of theory. Inclusion of electron correlation effects was performed from calculations at the MP2 and MP4(SDQ) levels. The frozen-core approximation was used to reduce the expensiveness of the calculations. Nevertheless, the reliability of this approach was verified by comparison with results determined from MP4 computations performed without the frozen-core approximation, as well as from CIPSI calculations. For simplicity, the following MP4 calculations with the frozen-core approximation are denoted simply as MP4, while those considering contribution from inner electrons are indicated by MP4(full). In all cases the 6-31G* basis set was used.

DFT calculations were carried out using different functionals at the local and nonlocal levels. Two local functionals were considered, which share the Slater term for exchange[47] but differ in the evaluation of correlation energy. This term was determined using the Vosko–Wilk–Nusair[48] and the Perdew-81[49] functionals. Following the nomenclature in the computer program Gaussian 94,[50] the two local functionals are referred to here as SVWN and SPL, respectively. One pure nonlocal DFT method was defined by combination of the Becke 88[51] and Perdew 86[52] functionals for exchange and correlation terms, respectively. This functional is denoted BP86. Finally, three hybrid nonlocal DFT methods were used, which share the Becke's three parameter functional[53] for exchange, while correlation was treated using Perdew 86, Lee–Yang–Parr,[54,55] and Perdew–Wang[56] functionals. The acronyms used in the following are B3P86, B3LYP, and B3PW91, respectively. In all cases the 6-31G* basis set was used for comparison with results derived from HF, MPx, and CIPSI calculations.

The MEP was computed using the standard procedure for all the wave functions. The MEP minima were located using a gradient minimization procedure until convergence in the depth of the MEP minimum was less than 0.01 kcal/mol. ESP derived charges were determined using Momany's strategy[57] and the optimization protocol described in detail elsewhere.[35,41,58]

All the wave functions were determined with Gaussian 94. MEPs were computed with either Gaussian 94 or the MOPETE/MOPFIT[59] computer program. ESP charges were calculated with MOPETE/MOPFIT. All the calculations were carried out on the IBM-SP2 computer of the Centre de Supercomputació de Catalunya (CESCA), as well as on SGI and HP workstations in our laboratory.

Results and Discussion

To facilitate discussion of the results, the comparative analysis of the MEP and MEP-derived properties is presented in three separate sections that correspond to: the ESP charges, as well as ESP and QM dipole moments; the MEP minima; and the MEP determined at different layers around the molecule. The values of these parameters computed at the different levels of theory for the series of molecules are not given for simplicity, but they are available upon request from the authors. In all cases the suitability of the DFT methods was examined by comparison with the results derived from MP4 calculations, which were used for reference purposes. At this point, the reliability of this calculational level was assessed to verify the goodness of the frozen-core approximation and the ability of MP4 results to reproduce the values determined at a higher level of electron correlation. This is briefly discussed at the beginning of each section.

ELECTROSTATIC CHARGES AND DIPOLES

Comparison of ESP charges, as well as ESP and exact QM dipole moments, determined from MP4 and MP4(full) calculations for selected molecules, permits us to examine the reliability of the frozen-core approximation. The results in Table I demonstrate that freezing of inner electrons introduces negligible changes in the magnitude of charges and dipoles even for formamide, for which correlation effects are very important (see below). There-

TABLE I.
ESP-Derived Charges and Dipole / QM Dipole Moments.

Molecule	Atom	Q	μ_{ESP}	μ_{QM}
H_2O	O	−0.798	2.198	2.145
		−0.796	*2.192*	*2.146*
	H	0.399		
		0.398		
CH_2O	C	0.430	2.216	2.219
		0.439	2.222	2.219
	H	−0.014		
		−0.016		
	O	−0.402		
		−0.407		
$HCONH_2$	N	−0.920	3.742	3.755
		−0.919	*3.742*	*3.755*
	H	0.419		
		0.418		
	H	0.392		
		0.392		
	C	0.648		
		0.648		
	O	−0.523		
		−0.523		
	H	−0.016		
		−0.016		

Electrostatic potential-derived charges (Q) are in electron units. Dipole moments (μ_{ESP}) and exact quantum mechanical dipole moments (μ_{QM}) are in debyes. Values were determined at the MP4 level with and without (values in italics) frozen-core approximation. See text for details.

fore, these results support the use of the frozen-core approximation for determination of these parameters. Furthermore, the ability of MP4 charges and dipoles to reproduce full configuration interaction (CI) values cannot be directly assessed owing to the difficulties in determining full-CI wave functions with the 6−31G* basis for the molecules studied here. Fortunately we demonstrated[60] that the full-CI values are reproduced nearly exactly by a CIPSI(G + M)* calculation, which allows us to use CIPSI(G + M) results to assess the performance of MP4 computations (see ref. 60 for full details). Figure 1 compares MP4 and CIPSI(G + M) ESP charges for H_2O, HF, CO, HCN, and CH_2O and shows that MP4 values correlate nearly perfectly with CIPSI(G + M) results. The scaling coefficient indicates a systematic deviation less than 1%. Furthermore, the root mean square (RMS) deviation is very small (0.017 electron units) and lies within the expected statistical error in the charge fitting. Comparison of ESP and QM dipole moments provides very similar results, which confirms the quality of these electrostatic properties determined at the MP4 level.

The preceding results validate the use of MP4 charges and dipole moments as reference data to analyze the suitability of the DFT methods. Indeed, for comparison purposes the study was extended to the values estimated from HF and MP2 calculations. ESP charges and dipoles, as well as QM dipole moments, were determined at the different levels of theory for 27 prototypical neutral molecules: HF, H_2O, HCN, NH_3, CO, CO_2, CH_2O, CH_3OH, CH_3NH_2, CF_2O, C_2H_4, C_2H_2, CH_3COH, $HCONH_2$, CHNHOH, $FCONH_2$, $O(CH_3)_2$, $CO(CH_3)_2$, CH_3CONH_2, $CO(NH_2)_2$, $N(CH_3)_3$, benzene, imidazole, pyrrole, furan, pyridine, and pyrimidine. In addition, a few ions were also considered to examine the performance of DFT methods for charged molecules: H_3O^+, NH_4^+, CH_3OH^+, pyridinium cation, OH^-, NH_2^-, $HCOO^-$, and phenolate anion.

* In the CIPSI method a space of generator determinants {G} is defined, which bring forth all single and double excitations. Those generated determinants {GD} contributing significantly to the first-order wave function are added to {G}, which is iteratively improved. At the end of this process, those determinants of {GD} contributing to the first-order wave function by a coefficient higher than a given threshold define the space {M}. Then, the correlation energy is estimated from two contributions: a variational term obtained by diagonalizing the Hamiltonian matrix representation in the space {G} + {M}, and a perturbation contribution up to the second-order including the {GD} − {M} determinants.

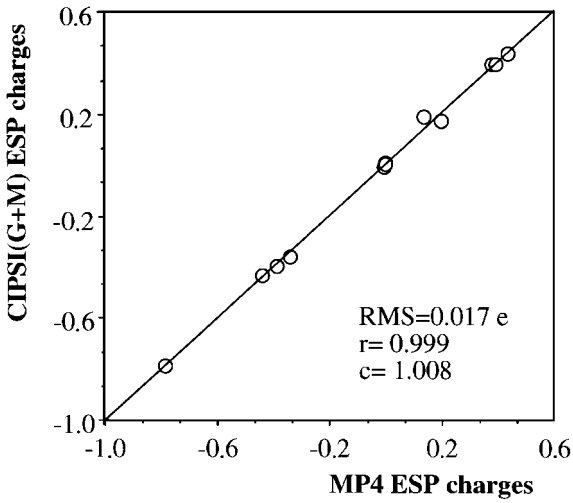

FIGURE 1. Correlation plot between ESP charges (in units of electron) determined from CIPSI(G + M) and MP4 calculations. See text for details.

Results of the statistical analysis for ESP charges of neutral molecules are in Table II. Reasonable agreement is found for all the methods, even at the HF level, as noted previously for a smaller series of molecules.[60] In fact, HF ESP charges deviate, on average, only 0.036 electron units from MP4 values due to the systematic enhancement (around 6%) of molecular polarity, which stems from underestimation of electron correlation effects at the HF level. Local (SVWN and SPL) functionals exaggerate these effects because the molecular charge separation is underestimated. The performance of these methods is slightly worse (RMS errors near 0.05 electron units) than that of HF, which warns against their use to determine ESP charges.[61] Simi-

TABLE II.
Results of Statistical Analysis of ESP-Derived Charges for Neutral Molecules.

Method	r[a]	c[b]	RMS
HF	0.998	0.942	0.036
SVWN	0.994	1.060	0.048
SPL	0.995	1.059	0.047
BP86	0.998	1.077	0.041
B3P86	0.997	1.039	0.033
B3LYP	0.999	1.058	0.028
B3PW91	0.997	1.040	0.033
MP2	0.998	1.012	0.020

Results were determined at different levels of theory with regard to the values derived from MP4 calculations. See text for details. RMS, the root mean square deviation (in electron units).
[a] The Pearson correlation coefficient.
[b] The optimum (MP4 / other method) ratio.

lar trends are found for the nonlocal BP86 functional, because the estimated charges are smaller (around 8%) than MP4 values. Finally, the best results are those derived from hybrid nonlocal DFT methods, which underestimate the molecular polarity by 4–6%, with the RMS deviations being around 0.03 electron units. In general the three hybrid functionals provide values very similar to the MP4 results. This suggests that the exchange part of the functional is more important than the correlation one for determination of electrostatic properties, in agreement with previous studies.[62, 63] However, MP2 ESP charges are the most accurate, because the RMS error is only 0.02 electron units from the MP4 values and the systematic deviation of polarity amounts to 1%.

All the methods provide reasonably good ESP charges for cations and anions (see Table III). HF charges deviate by around 7% from MP4 values, a deviation similar to that found for neutral molecules. All the DFT methods perform better than HF. They exaggerate the influence of electron correlation effects, which leads to underestimation of the charge separation ranging from 1.5% (B3P86) to 4.7% (BP86). The RMS error of MP2 charges is slightly larger than that found for hybrid nonlocal DFT methods, but the systematic deviation of polarity in DFT calculations is corrected at this level of theory as noted by the MP4/MP2 ratio (c = 1.005). When cations and anions are analyzed separately, the statistical results for anions are worse than for cations. Nevertheless, the performance of all the methods follows the same trends mentioned above.

TABLE III.
Results of Statistical Analysis of ESP-Derived Charges for Cations and Anions.

Method	r[a]	c[b]	RMS
HF	0.995	0.926	0.056
SVWN	0.997	1.025	0.037
SPL	0.997	1.028	0.037
BP86	0.998	1.047	0.032
B3P86	0.999	1.015	0.018
B3LYP	0.999	1.031	0.019
B3PW91	0.999	1.016	0.017
MP2	0.998	1.005	0.031

Results were determined at different levels of theory with regard to the values derived from MP4 calculations. See text for details. RMS, the root mean square deviation (in electron units).
[a] The Pearson correlation coefficient
[b] The optimum (MP4 / other method) ratio.

Additional information on the quality of the electrostatic distribution is gained from inspection of dipole moments. Both ESP and exact QM dipoles were examined for neutral molecules with nonvanishing dipole moments ($\mu \neq 0$). The statistical analysis is given in Table IV. As expected, almost identical results are obtained for QM and ESP dipoles, which supports the reliability of the fitting procedure between QM and classical MEPs. The dipoles determined from the different methods show reasonably good agreement with MP4 values, which indicates that electron correlation effects are not extremely important for this property. The largest deviations are found for the HF results (see Table IV), which overestimates the dipoles by 7–8% [RMS error of 0.24 D (debye)]. All the DFT methods underestimate the dipole moment, which indicates an overestimation of correlation effects. The deviation amounts to 2% for the local functionals and 4% for the pure nonlocal BP86. Hybrid nonlocal methods provide the best dipole, which are underestimated by around 1% (RMS errors of 0.05–0.08 D). As noted for the charges, the B3P86 results are especially good. However, the best agreement is found at the MP2 level, because the dipoles are underestimated by only 0.2%.

TABLE IV.
Results of Statistical Analysis of ESP-Derived and Exact QM Dipole Moments for Neutral Molecules.

Method	r[a]	c[b]	RMS
HF	0.992	0.924	0.238
	0.993	*0.928*	*0.233*
SVWN	0.996	1.018	0.099
	0.992	*1.021*	*0.105*
SPL	0.995	1.022	0.108
	0.996	*1.024*	*0.109*
BP86	0.995	1.039	0.134
	0.995	*1.042*	*0.136*
B3P86	0.998	1.003	0.055
	0.998	*1.006*	*0.055*
B3LYP	0.998	1.016	0.074
	0.996	*1.018*	*0.076*
B3PW91	0.998	1.007	0.061
	0.998	*1.010*	*0.061*
MP2	0.998	1.002	0.047
	0.998	*1.003*	*0.047*

Results were determined at different levels of theory with regard to the values derived from MP4 calculations. The quantum mechanical (QM) values are in italic. See text for details. RMS, the root mean square deviation (in debyes).
[a] The Pearson correlation coefficient
[b] The optimum (MP4 / other method) ratio.

It is worth noting that all the DFT methods provide dipole moments more similar to the MP4 values than the HF ones. This finding is surprising according to the statistical results for comparison of charges. Thus, errors in DFT charges, which are even larger than the deviation of the HF values, are not clearly reflected in the errors of the corresponding dipole moments (see Tables II, IV). This most likely stems from the simplified description of the molecular charge distribution in terms of the dipole moment and argues against the exclusive use of dipole moments to examine the overall quality of the electrostatic distribution.

Further insight is provided by comparison of theoretical and gas phase experimental dipole moments. Results in Table V show that reasonable agreement is generally found with the experimental values, the average RMS deviation being around 0.2 D. The best agreement is found for MP*x* results, while local and nonlocal BP86 functionals show the largest deviations. Although the HF dipoles are overestimated by 8%, all the DFT methods underestimate the experimental dipole. As noted before, results from hybrid nonlocal calculations, particularly the B3P86 functional, exhibit the closest agreement with experimental values. All these findings reflect the trends mentioned previously for comparison with regard to the MP4 results. Inspection of the results suggests that correlation effects at the MP*x* level corrects, on average, the systematic overestimation found in HF results. Nevertheless, MP*x* calculations have not completely eliminated the random error in the results. This most likely originates from the use of a small basis rather than from the treatment of correlation effects.

In summary, the statistical analysis of charges for neutral molecules permits the establishment of the following ordering of quality: MP2 > B3P86 ≥ B3PW91 ≈ B3LYP > HF > BP86 > SVWN ≈ SPL. For charged molecules such an ordering is MP2 ≈ B3P86 ≈ B3PW91 ≥ B3LYP > BP86 ≈ SPL ≈ SVWN > HF. From these results only hybrid nonlocal DFT methods are expected to provide better ESP charges than the HF method. With respect to the dipole moments the following ordering can be defined using the MP4 values as reference data: MP2 > B3P86 > B3PW91 > B3LYP > SVWN > SPL > BP86 > HF. This well reflects the ordering established when the experimental dipoles are considered as reference: MP4 ≈ MP2 ≈ B3P86 ≈ B3PW91 ≈ B3LYP > SVWN > SPL > BP86 > HF.

MEP MINIMA

The suitability of the frozen-core approximation for estimation of MEP minima was also examined by comparison between MP4 and MP4(full) results for a limited series of small molecules (H_2O, CH_2O, NH_3, CO_2, CH_2CH_2, and $HCONH_2$). The results (data not shown, but available upon request) demonstrate that such approximation has no effect on the depth of the MEP minimum and supports its use in calculations. Furthermore, comparison between MP4 and CIPSI(G + M) MEP minima (Fig. 2) for a limited series of small

TABLE V.
Results of Statistical Analysis of ESP-Derived Dipole Moments.

Method	*r* [a]	*c* [b]	RMS
HF	0.990	0.922	0.211
SVWN	0.977	1.024	0.211
SPL	0.976	1.030	0.217
BP86	0.977	1.050	0.226
B3P86	0.985	1.005	0.169
B3LYP	0.985	1.020	0.171
B3PW91	0.986	1.010	0.166
MP2	0.980	1.000	0.191
MP4	0.985	0.996	0.164

Results were determined at different levels of theory with regard to the gas phase experimental values for neutral molecules with nonzero dipole. See text for details. RMS, the root mean square deviation (in debyes).
[a] The Pearson correlation coefficient.
[b] The optimum (experimental / theoretical) ratio.

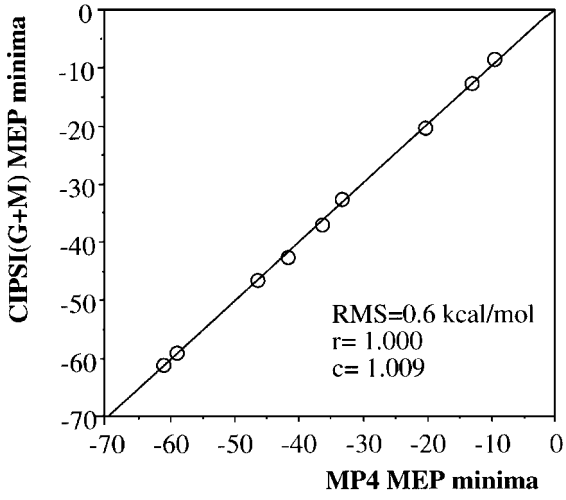

FIGURE 2. Comparison of the depth of MEP minima (in kcal / mol) determined from CIPSI(G + M) and MP4 calculations. See text for details.

molecules (H_2O, HF, CO, HCN, and CH_2O) demonstrate the reliability of MP4 values, as noted in the small RMS deviation (0.6 kcal/mol), the excellent Pearson correlation coefficient ($r = 1.000$), and the value of the scaling coefficient ($c = 1.009$), which indicates a systematic deviation of only 0.9%.

A total set of 40 MEP minima were determined for the neutral molecules H_2, N_2, H_2O, NH_3, HF, F_2, CO, CO_2, HCN, CH_2O, CH_3OH, CH_3NH_2, CH_3COH, CHCH, CH_2CH_2, COF_2, $CO(CH_3)_2$, CH_3CONH_2, $CHONH_2$, $COFNH_2$, $CO(NH_2)_2$, CHNHOH, $O(CH_3)_2$, $N(CH_3)_3$, pyrrole, pyridine, imidazole, pyrimidine, and furane. The analysis was not extended to charged molecules, because information about the depth and distribution of MEP minima is especially useful for neutral compounds. Table VI gives the results of the statistical analysis for comparison with MEP minima determined at the MP4 level.

As expected from previous studies,[60] all the methods reproduce the depth of MEP minima reasonably well. The HF results are moderately deviated from MP4 values (4.6 kcal/mol), and the systematic overestimation (only 8%) agrees with the enhancement of polarity discussed before for charges and dipoles. The DFT methods provide better statistical results than HF. The depth of MEP minima is underestimated by a factor ranging from 7% (BP86) to 2% (B3P86), which reflects the lower polarity arising from overestimation of electron correlation effects. Hybrid nonlocal methods, especially B3P86, provide the best results, because the RMS error is around 1 kcal/mol. Finally, MP2 calculations correct the systematic underestimation of the depth of MEP minima in DFT calculations, as stated by the scaling coefficient ($c = 1.009$). Overall, the following ordering can be defined for the different methods: MP2 ≈ B3P86 ≈ B3PW91 > B3LYP > SPL > BP86 ≈ SVWN > HF. These results suggest that inclusion of gradient corrections, and particularly treatment of exact exchange, should improve the description provided by local functionals of those molecular properties that can be related to the depth of MEP minima,[4–7] such as formation of hydrogen bonds or proton affinities.[53, 64, 65]

Deeper insight is gained upon comparison of the MEP minima in different classes: I, MEP minima in small molecules (H_2, N_2, F_2, HCN, CO); II, MEP minima of carbonyl derivatives [CO_2, CH_2O, CH_3COH, $CO(CH_3)_2$, CH_3CONH_2, CF_2O, $COFNH_2$, $CO(NH_2)_2$]; III, MEP minima of derivatives of water and ammonia [H_2O, NH_3, CH_3OH, CH_3NH_2, CHNHOH, $O(CH_2)_3$, $N(CH_3)_3$]; IV, MEP minima due to lone pairs of aromatic compounds (pyridine, pyrimidine, furane, imidazole); and V, MEP minima related to π-charge distribution (CH_2CH_2, CHCH, benzene, pyridine, pyrimidine, furane, imidazole, pyrrole). Results of the statistical analysis are shown in Table VII.

The RMS errors for MEP minima of small molecules (class I) ranges from 1.2 to 2.5 kcal/mol. The largest absolute discrepancies are found at the HF level, while the greatest relative deviations occur for local and nonlocal BP86 functionals, which underestimate the depth of MEP minima by 6–7%. The best agreement with MP4 values is found for MP2 and hybrid nonlocal results. Similar trends occur in classes II–IV, so that we limit the discussion to some relevant features. In class II the errors at the HF, local, and BP86 levels are notably enlarged; the reverse effect is observed for hybrid nonlocal functionals. Local and nonlocal BP86 results are worse than HF values in class III, and hybrid nonlocal results are worse than MP2 ones. Finally, the behavior of local and BP86 functionals is poor in class IV, because the RMS errors are near 6 kcal/mol. and the depth of MEP minima is underestimated by 10%. Hybrid nonlocal methods improve the results, and RMS errors around 1.4 kcal/mol are found for B3PW91 and MP2. The results of the statistical analysis for MEP minima due to π-charge distribution (class V) reveal a

TABLE VI.
Results of Statistical Analysis of MEP Minima for Neutral Molecules.

Method	r [a]	c [b]	RMS
HF	0.994	0.923	4.6
SVWN	0.996	1.054	3.1
SPL	0.995	1.062	2.0
BP86	0.996	1.071	2.6
B3P86	0.998	1.021	0.7
B3LYP	0.998	1.033	1.3
B3PW91	0.998	1.025	0.8
MP2	0.999	1.009	1.0

Results were determined at different levels of theory with regard to the values derived from MP4 calculations. See text for details. RMS, the root mean square deviation (in kcal/mol).

[a] The Pearson correlation coefficient.

[b] The optimum (MP4/other method) ratio.

TABLE VII.
Results of Statistical Analysis of MEP Minima.

Class / Parameter	HF	SVWN	SPL	BP86	B3P86	B3LYP	B3PW91	MP2
I								
r[a]	0.987	0.994	0.994	0.994	0.998	0.998	0.998	0.996
c[b]	0.987	1.058	1.059	1.068	1.045	1.040	1.042	0.981
RMS	2.5	1.9	2.0	2.1	1.2	1.2	1.3	1.4
II								
r	0.998	0.988	0.997	0.996	0.999	0.998	0.999	0.999
c	0.987	1.057	1.070	1.058	1.007	1.015	1.009	1.027
RMS	6.9	2.7	3.6	2.7	0.6	0.9	0.7	1.2
III								
r	0.994	0.988	0.986	0.987	0.996	0.994	0.996	0.999
c	0.956	1.043	1.048	1.064	1.025	1.040	1.025	1.000
RMS	3.4	3.8	3.5	4.8	2.2	3.1	1.9	0.6
IV								
r	0.995	0.995	0.995	0.995	0.998	0.998	0.998	0.999
c	0.939	1.100	1.104	1.104	1.051	1.056	1.020	1.000
RMS	4.2	5.7	6.0	5.9	3.1	3.4	1.4	1.5
V								
r	0.997	0.997	0.997	0.997	0.998	0.992	0.998	0.999
c	0.890	0.947	0.956	0.936	0.938	1.012	0.935	0.994
RMS	2.3	1.1	0.9	0.5	1.1	0.7	1.2	0.2

Results were determined at different levels of theory with regard to the values derived from MP4 calculations for different types of minima. See text for details. RMS, the root mean square deviation (in kcal / mol).
[a] The Pearson correlation coefficient.
[b] The optimum (MP4 / other method) ratio.

different behavior. Thus, the HF method, which has the largest error (RMS of 2.3 kcal/mol), overestimates the depth of MEP minima. With the exception of B3LYP, all the DFT methods give similar results: the RMS errors range from 0.5 to 1.2 kcal/mol and the depth of MEP minima is overestimated by 4–6%. These results are still worse than those found for the MP2 method, which shows the best agreement with MP4 results (RMS of 0.2 kcal/mol).

In summary, the worse results are generally found at the HF, local, and pure nonlocal BP86 levels. The MP2 values show the best agreement. Hybrid nonlocal methods perform well for classes I and II, but the results are slightly poorer than the MP2 ones found for classes III and IV, especially for minima originated from lone pairs of aromatic compounds, DFT methods underestimate the depth of MEP minima for classes I–IV, which agrees with the findings discussed previously for charges and dipoles. Nevertheless, such an effect is reversed for MEP minima in class V. This suggests that introduction of electron correlation effects by DFT methods does not occur at the same extent in all the space surrounding the atoms.

MEP DISTRIBUTION IN LAYERS

The quality of the electrostatic distribution was also analyzed by examining the MEP computed at eight layers around the molecule. The layers were defined as molecular surfaces determined by scaling of the van der Waals radius of atoms by a factor (λ) of 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, and 2.0, which allows us to compare the MEP in regions near and outside the van der Waals surface. It is precisely in these outer regions where the MEP dominates the intermolecular interactions. This is helpful for the potential application of the MEP determined from DFT methods to study noncovalent interactions, as well as to examine their reliability in the framework of self-consistent reaction field (SCRF) continuum models.[25, 27, 66]

To carry out this analysis, it is reasonable to assume from the preceding sections that MP4(full)

results reproduce the exact full-CI MEP in these regions. Indeed, inspection of Figure 3 shows that the frozen-core approximation has a negligible effect on the MEP determined at these layers, because the RMS error between the MP4 and MP4(full) MEPs for H_2O, CH_2O, and $HCONH_2$ is less than 0.04 kcal/mol. Therefore, the MEP computed in those layers from the various methods was compared with the MP4 MEP for the series of neutral molecules. Table VIII gives the average RMS deviations at each layer, and selected RMS plots are shown in Figure 4.

In all cases the RMS deviation decreases monotonously as the distance of the layer from the molecule increases. This is not surprising because

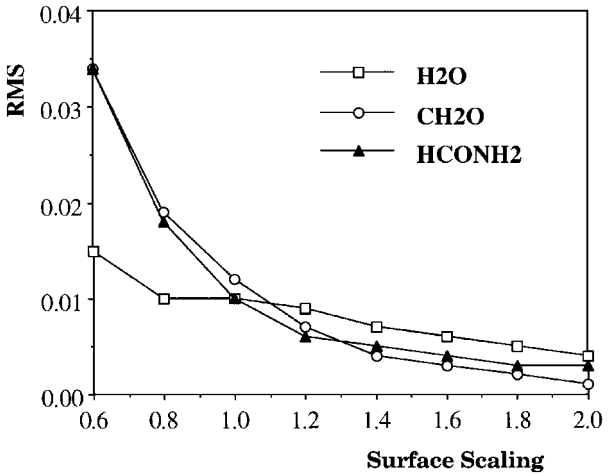

FIGURE 3. Root mean square (RMS) errors (in kcal/mol) between the MEPs computed at the MP4 level with and without frozen-core approximation. The MEPs were determined for water, formaldehyde, and formamide at the different layers ranging from 0.6 to 2.0 times the van der Waals radius. See text for details.

electron correlation effects are relevant mainly in those regions of space closest to the nuclei.[60] The best agreement is found for the MP2 method, for which the average RMS errors decrease from 1.1 ($\lambda = 0.6$) to 0.1 ($\lambda = 2.0$) kcal/mol. The largest errors are found for HF, local, and nonlocal BP86 methods, the deviation ranging from around 5.0 ($\lambda = 0.6$) to 0.5 ($\lambda = 2.0$) kcal/mol. The best DFT results are provided by hybrid nonlocal methods, the agreement being slightly better for B3LYP. Thus, at the van der Waals surface ($\lambda = 1.0$) the average RMS error ranges from 1.7 to 2.0 kcal/mol at the HF local DFT levels, while it amounts to 0.9 kcal/mol for B3LYP and only 0.5 kcal/mol for MP2. It is also interesting to note the results for the layer at $\lambda = 1.2$, because the resulting molecular surface is used to simulate the solute/solvent interface in various SCRF methods.[67–72] Results in Table VIII show that the HF MEP, as well as that computed from local and BP86 functionals, deviate on average more than 1 kcal/mol from the MP4 MEP. The error is 0.7 kcal/mol for hydrid nonlocal DFT methods and decreases to 0.4 kcal/mol at the MP2 level. Accordingly, calculation of the solvent RF at the MP2 and hybrid nonlocal DFT levels is expected to be highly accurate. Indeed, the results suggest that parametrization of SCRF continuum models at the HF or local DFT levels is probably not directly transferable to hybrid nonlocal DFT and MP2 methods.

In spite of the general trends mentioned above, there are relevant differences in the RMS profiles of the various methods for diverse molecules. This is illustrated in Figure 4, which shows the RMS profiles for water, methanol, formamide, and pyridine. Thus, in the case of water the largest deviations at layers $\lambda \geq 0.8$ are found for BP86 and B3LYP, the errors being even larger than those

TABLE VIII.
Average RMSD Between MEPs Computed at Different Levels of Theory.

Layer	HF	SVWN	SPL	BP86	B3P86	B3LYP	B3PW91	MP2
0.6	4.7	5.0	4.6	4.2	4.0	3.2	4.2	1.1
0.8	2.9	2.9	2.7	2.4	2.0	1.5	2.1	0.8
1.0	2.0	1.7	1.8	1.5	1.1	0.9	1.2	0.5
1.2	1.5	1.1	1.1	1.1	0.7	0.7	0.7	0.4
1.4	1.2	0.8	0.8	0.8	0.5	0.5	0.5	0.3
1.6	1.0	0.6	0.6	0.6	0.3	0.4	0.4	0.2
1.8	0.7	0.5	0.5	0.5	0.3	0.3	0.3	0.2
2.0	0.6	0.4	0.4	0.4	0.2	0.3	0.2	0.1

RMSD, root mean square deviations (in kcal/mol), computations were with regard to the values derived from MP4 calculations at selected layers from the molecules. See text for details.

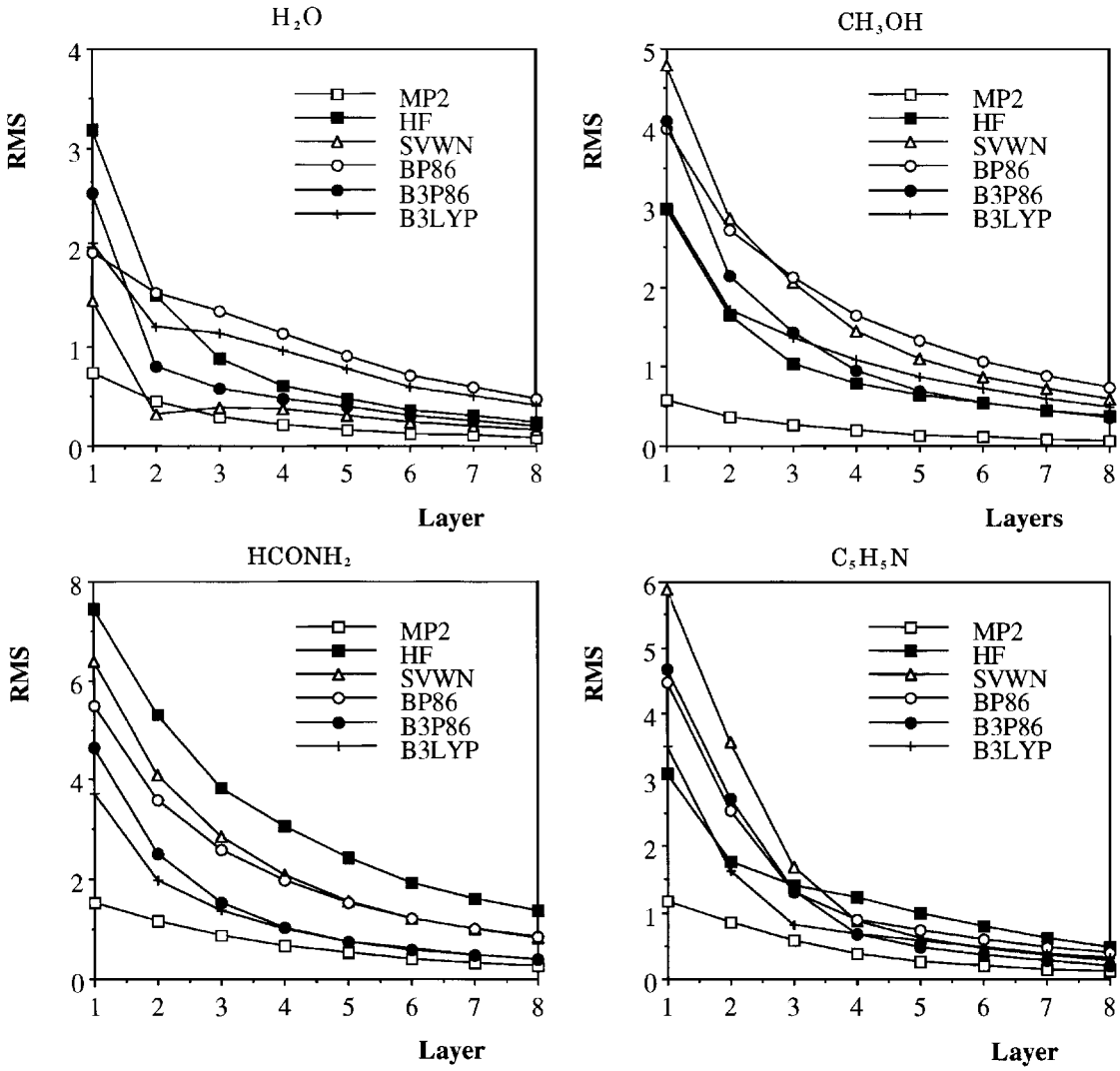

FIGURE 4. Root mean square (RMS) errors (in kcal / mol) between the MEPs computed at the HF, SVWN, BP86, B3P86, B3LYP, and MP2 levels with regard to the MP4 values. The plots correspond to water, methanol, formamide, and pyridine. The RMS errors are given for the eight layers ranging from 0.6 to 2.0 times the van der Waals radius. The profiles for SPL and B3PW91 (not given) are very similar to those of SVWN and B3P86, respectively. See text for details.

found for the HF MEP, whereas the MEP determined from local DFT methods is close to the MP4 MEP. However, local and nonlocal BP86 functionals for methanol provide the worst results, and the closest agreement with MP4 values, in addition to the MP2 method, is obtained for hybrid nonlocal DFT and HF methods. The RMS profiles for formamide clearly reflect the trends noted from the results in Table VIII: the agreement with the MP4 values are MP2 > hybrid nonlocal > BP86 ≥ local > HF. Finally, for pyridine the largest deviations are found for local functionals at layers $\lambda \leq 1.0$, and for the HF method at layers $\lambda \geq 1.2$. Indeed, the deviations for local and nonlocal DFT methods are very similar at layers $\lambda \geq 1.2$. Therefore, as noted before, the present results suggest that inclusion of electron correlation effects by DFT methods is not uniform in all the regions of space surrounding the molecule and the magnitude of such correlation effects shows strong fluctuations for different molecules.

Conclusion

The comparative analysis of the MEP and MEP-derived properties allows us to extract some general conclusions about the performance of DFT

methods for calculating electrostatic properties. The results clearly show the well-known tendency of the HF method to overestimate the molecular charge separation due to neglect of electron correlation effects. This effect is greatly corrected by inclusion of electron correlation at the MP2 level, whose results nearly reproduce the electrostatic properties (charges, dipoles, MEP minima, and the MEP distribution around the molecule) determined from the more expensive MP4 calculations, even though they show a slight, but systematic, underestimation.

All the DFT methods exaggerate the magnitude of electron correlation effects, because the values for the diverse electrostatic properties are clearly underestimated with regard to the MP4 values. Further, this trend is reflected in the comparison of the calculated and experimental results. Nevertheless, the magnitude of such underestimation shows notable differences between the various functionals considered in the study. At this point, the results state that hybrid nonlocal functionals perform better sensibly than simple gradient-corrected and local functionals. This is consistent for all the electrostatic properties included in the analysis with MP4 results, but also in the inspection of dipoles moments in front of the experimental values. This indeed indicates that the exchange part of the functional is more important than the correlation one for determination of molecular electrostatic properties. The statistical analysis indicates that B3P86 appears to provide slightly better results than B3LYP and B3PW91, but the differences are very small and they are unlikely to be significant.

The comparison analysis of the MEP and MEP-derived properties has implications in the application of DFT methods to noncovalent interactions. Very recently Reynolds et al. also examined the suitability of nonlocal DFT in the parametrization of force-field methods,[73] with particular attention to the development of strategies for including explicit polarization and higher order electrostatic effects. Present results indicate that local DFT, or even DFT with simple inclusion of gradient corrections, are not well suited to the study of hydrogen-bond systems or ionic clusters, which are very important in biological and pharmacological systems. On the contrary, the mixing of an exact exchange term with DFT exchange and correlation makes hybrid nonlocal functionals a promising computational strategy for the study of such biomolecular systems. These methods could give reliable results for biomolecules with an accuracy comparable to that of MP2 calculations, but at a sensibly lower computational cost.

Acknowledgments

Prof. C. A. Reynolds is kindly acknowledged for encouraging discussion and for a preprint of his manuscript. We thank the Centre de Supercomputació de Catalunya (CESCA, Mol. Recog. Project) for computational facilities and the Dirección General de Investigación Científica y Técnica (DGICYT, Grants PB93-0779 and PB94-0940) for financial support. R. S. is the recipient of a predoctoral grant from the Comissió Interministerial de Recerca i Innovació Tecnològica.

References

1. E. Scrocco and J. Tomasi, *J. Comput. Chem.*, **42,** 95 (1973).
2. E. Scrocco and J. Tomasi, *Adv. Quantum Chem.*, **11,** 115 (1978).
3. P. Politzer and K. C. Daiker, In *The Force Concept in Chemistry*, D. M. Debb, Ed., Van Nostrand Reinhold, Amsterdam, 1981, p. 294.
4. P. Politzer and D. G. Truhlar, Eds., *Chemical Applications of Atomic and Molecular Electrostatic Potentials*, Plenum, New York, 1981.
5. P. Politzer and J. S. Murray, In *Reviews in Computational Chemistry*, K. B. Lipkowitz and D. B. Boyd, Eds., VCH, New York, 1991, p. 273.
6. J. Tomasi, G. Alagona, R. Bonaccorsi, C. Ghio, and R. Cammi, In *Theoretical Models in Chemical Bonding, Part IV*, Z. B. Maksic, Ed., Springer–Verlag, Berlin, 1991, p. 229.
7. P. Kollman, J. McKelvey, A. Johansson, and S. Rothenberg, *J. Am. Chem. Soc.*, **97,** 955 (1975).
8. J. S. Murray, B. A. Zilles, K. Jayaruriya, and P. Politzer, *J. Am. Chem. Soc.*, **108,** 915 (1986).
9. M. Orozco and F. J. Luque, *J. Comput. Chem.*, **14,** 587 (1993).
10. G. Náray–Szabó and G. G. Ferenczy, *Chem. Rev.*, **95,** 829 (1995).
11. M. Orozco and F. J. Luque, In *Molecular Electrostatic Potentials: Concepts and Applications, Theoretical and Computational Chemistry, Vol. 3*, J. S. Murray and K. Sen, Eds., Elsevier, Amsterdam, 1996, p. 181.
12. J. Tomasi, B. Menucci, and R. Cammi, In *Molecular Electrostatic Potentials: Concepts and Applications, Theoretical and Computational Chemistry, Vol. 3*, J. S. Murray and K. Sen, Eds., Elsevier, Amsterdam, 1996, p. 1.
13. N. G. J. Richards and S. L. Price, *Int. J. Quantum Chem., Quantum Biol. Symp.*, **16,** 73 (1989).
14. C. Burt, W. G. Richards, and P. Huxley, *J. Comput. Chem.*, **11,** 1139 (1990).
15. A. M. Richard, *J. Comput. Chem.*, **12,** 959 (1991).

16. J. Rodríguez, F. Manaut, and F. Sanz, *J. Comput. Chem.*, **14,** 922 (1993).
17. J. D. Petke, *J. Comput. Chem.*, **14,** 928 (1993).
18. E. Besalú, R. Carbó, J. Mestres, and M. Solà, In *Topics in Current Chemistry: Molecular Similarity, Vol. 173*, K. Sen, Ed., Springer–Verlag, Berlin, 1995, p. 31.
19. C. Lecomte, N. Ghermani, V. Pichon–Pesme, and M. Souhasson, *J. Mol. Struct. (Theochem.)*, **255,** 241 (1992).
20. J. C. White and A. C. Hess, *J. Phys. Chem.*, **97,** 6398 (1993).
21. R. F. Stewart and B. M. Craven, *Biophys. J.*, **65,** 998 (1993).
22. M. K. Gilson and B. Honig, *Proteins*, **4,** 7 (1988).
23. R. J. Zauhar and R. S. Morgan, *J. Comput. Chem.*, **9,** 171 (1988).
24. P. Politzer, P. Lane, J. S. Murray, and T. Brinck, *J. Phys. Chem.*, **96,** 7938 (1992).
25. J. Tomasi and M. Persico, *Chem. Rev.*, **94,** 2027 (1994).
26. C. Alhambra, F. J. Luque, and M. Orozco, *J. Phys. Chem.*, **99,** 3084 (1995).
27. J. L. Rivail and D. Rinaldi, In *Computational Chemistry, Review of Current Trends*, J. Leszczynski, Ed., World Scientific Publishing, Singapore, 1996, p. 139.
28. K. S. Sen and P. Politzer, *J. Chem. Phys.*, **90,** 4370 (1989).
29. S. R. Gadre, S. A. Kulkarni, and I. H. Shrivastava, *J. Chem. Phys.*, **96,** 5253 (1992).
30. S. R. Gadre, S. A. Kulkarni, C. H. Suresh, and I. H. Shrivastava, *Chem. Phys. Lett.*, **239,** 273 (1995).
31. L. E. Chirlian and M. M. Francl, *J. Comput. Chem.*, **8,** 894 (1987).
32. D. E. Williams, *J. Comput. Chem.*, **9,** 745 (1988).
33. R. J. Woods, M. Khalil, W. Pell, S. H. Moffat, and V. H. Smith, *J. Comput. Chem.*, **11,** 297 (1990).
34. C. M. Breneman and K. B. Wiberg, *J. Comput. Chem.*, **11,** 361 (1990).
35. M. Orozco and F. J. Luque, *J. Comput. Chem.*, **11,** 909 (1990).
36. C. Alemán, F. J. Luque, and M. Orozco, *J. Computer Aided Mol. Des.*, **7,** 721 (1993).
37. W. D. Cornell, P. Cieplak, C. I. Bayly, and P. A. Kollman, *J. Am. Chem. Soc.*, **113**, 5203 (1991).
38. P. C. Hariharan and J. A. Pople, *Theor. Chim. Acta*, **28,** 213 (1973).
39. F. J. Luque, F. Illas and M. Orozco, *J. Comput. Chem.*, **11,** 416 (1990).
40. G. G. Ferenczy, C. A. Reynolds, and W. G. Richards, *J. Comput. Chem.*, **11,** 159 (1990).
41. B. H. Besler, K. M. Merz, Jr., and P. A. Kollman, *J. Comput. Chem.*, **11,** 431 (1990).
42. B. Wang and G. P. Ford, *J. Comput. Chem.*, **15,** 200 (1994).
43. C. Alhambra, F. J. Luque, and M. Orozco, *J. Comput. Chem.*, **15,** 12 (1994).
44. C. Møller and M. S. Plesset, *Phys. Rev.*, **46,** 618 (1934).
45. B. Huron, J. P. Malrieu, and P. Rancurel, *J. Chem. Phys.*, **58,** 5745 (1973).
46. M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. A. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head–Gordon, C. Gonzalez, and J. A. Pople, Eds., *Gaussian 94 (Rev. A.1)*, Gaussian, Inc., Pittsburgh, PA, 1995.
47. W. Kohn and L. J. Sham, *Phys. Rev.*, **140,** A1133 (1965).
48. S. H. Vosko, L. Wilk, and M. Nusair, *Can. J. Phys.*, **58,** 1200 (1980).
49. J. P. Perdew and A. Zunger, *Phys. Rev. B*, **23,** 5048 (1981).
50. M. J. Frisch, A. Frisch, and J. B. Foresman, *Gaussian 94 User's Reference*, Gaussian, Inc., Pittsburgh, PA, 1995.
51. A. D. Becke, *Phys. Rev. A*, **38,** 3098 (1988).
52. J. P. Perdew, *Phys. Rev. B*, **33,** 8822 (1986).
53. A. D. Becke, *J. Chem. Phys.*, **98,** 5648 (1993).
54. C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B*, **37,** 785 (1988).
55. B. Miehlich, A. Savin, H. Stoll, and H. Preuss, *Chem. Phys. Lett.*, **157,** 200 (1989).
56. J. P. Perdew and Y. Wang, *Phys. Rev. B*, **45,** 13244 (1992).
57. F. A. Momany, *J. Phys. Chem.*, **82,** 592 (1978).
58. M. Orozco and F. J. Luque, *J. Computer Aided Mol. Des.*, **4,** 411 (1990).
59. F. J. Luque and M. Orozco, Unpublished version of MOPETE/MOPFIT program, University of Barcelona, 1995.
60. F. J. Luque, M. Orozco, F. Illas, and J. Rubio, *J. Am. Chem. Soc.*, **113,** 5203 (1991).
61. A. St.-Amant, W. D. Cornell, P. A. Kollman, and T. A. Halgren, *J. Comput. Chem.*, **16,** 1483 (1995).
62. F. De Proft, J. M. L. Martin, and P. Geerlins, *Chem. Phys. Lett.*, **250,** 393 (1996).
63. J. E. Wampler, *J. Chem. Inf. Sci.*, **35,** 617 (1995).
64. F. Sim, A. St.-Amant, I. Papaiand, and D. R. Salahub, *J. Am. Chem. Soc.*, **114,** 4391 (1992).
65. P. Hobza, J. Sponer, and T. Reschel, *J. Comput. Chem.*, **16,** 1315 (1995).
66. C. J. Cramer and D. G. Truhlar, In *Reviews in Computational Chemistry, Vol. 6*, K. B. Lipkowitz and D. B. Boyd, Eds., VCH, New York, 1995, p. 4.
67. I. Tuñon, M. F. Ruiz–Lopez, D. Rinaldi, and J. Bertran, *J. Comput. Chem.*, **17,** 148 (1996).
68. V. Dillet, D. Rinaldi, and J. L. Rivail, *J. Phys. Chem.*, **98,** 5034 (1994).
69. A. Klamt and G. Schüürmann, *J. Chem. Soc. Perkin Trans. II*, 799 (1993).
70. T. N. Truong and E. V. Stefanovich, *Chem. Phys. Lett.*, **240,** 253 (1990).
71. M. Orozco, M. Bachs, and F. J. Luque, *J. Comput. Chem.*, **16,** 563 (1995).
72. R. Bonaccorsi, F. Floris, and J. Tomasi, *J. Mol. Liquids*, **47,** 25 (1990).
73. P. J. Winn, G. G. Ferenczy, and C. A. Reynolds, *J. Phys. Chem.*, to appear.